

UC Berkeley

UC Berkeley Previously Published Works

Title

Molecular Mechanics Simulations and Improved Tight-Binding Hamiltonians for Artificial Light Harvesting Systems: Predicting Geometric Distributions, Disorder, and Spectroscopy of Chromophores in a Protein Environment.

Permalink

<https://escholarship.org/uc/item/7t87p05h>

Journal

The journal of physical chemistry. B, 122(51)

ISSN

1520-6106

Authors

Lee, Joonho
Lee, Donghyun
Kocherzhenko, Aleksey A
et al.

Publication Date

2018-12-01

DOI

10.1021/acs.jpcb.8b08858

Peer reviewed

Molecular Mechanics Simulations and Improved Tight-Binding Hamiltonians for Artificial Light Harvesting Systems: Predicting Geometric Distributions, Disorder, and Spectroscopy of Chromophores in a Protein Environment

Joonho Lee,^{†,‡} Donghyun Lee,[¶] Aleksey A. Kocherzhenko,^{¶,§} Loren Greenman,^{¶,||}

Daniel T. Finley,[¶] Matthew B. Francis,[¶] and K. Birgitta Whaley*,[¶]

[†]*Department of Chemistry, Pohang University of Science and Technology, Pohang,
Gyeongsangbuk-do, Korea*

[‡]*Present Address: College of Chemistry, University of California, Berkeley, California,
USA*

[¶]*College of Chemistry, University of California, Berkeley, California, USA*

[§]*Present Address: Department of Chemistry, California State University, East Bay, USA,
USA*

^{||}*Present Address: Department of Physics, Kansas State University, 116 Cardwell Hall,
1228 N. 17th St. Manhattan, Kansas, USA*

E-mail: whaley@berkeley.edu

Abstract

We present molecular mechanics and spectroscopic calculations on prototype artificial light harvesting systems consisting of chromophores attached to a tobacco mosaic virus (TMV) protein scaffold. These systems have been synthesized and characterized spectroscopically, but information about the microscopic configurations and geometry of these TMV-templated chromophore assemblies is largely unknown. We use a Monte Carlo conformational search algorithm to determine the preferred positions and orientations of two chromophores, Coumarin 343 together with its linker, and Oregon Green 488, when these are attached at two different sites (104 and 123) on the TMV protein. The resulting geometric information shows that the extent of disorder and aggregation properties, and therefore the optical properties of the TMV-templated chromophore assembly, are highly dependent on the choice of chromophores and protein site to which they are bound. We used the results of the conformational search as geometric parameters together with an improved tight-binding Hamiltonian to simulate the linear absorption spectra and compare with experimental spectral measurements. The ideal dipole approximation to the Hamiltonian is not valid since the distance between chromophores can be very small. We found that using the geometries from the conformational search is necessary to reproduce the features of the experimental spectral peaks.

1 Introduction

Light harvesting antennae of photosynthetic organisms are exquisitely organized biomolecular structures.^{1,2} Although nearly every photosynthetic species on the planet has evolved a light harvesting antenna that is customized to its environment, all these antennae are actually composed of relatively few types of pigment molecules, e.g., chlorophylls, bacteriochlorophylls, carotenoids, phycobiliproteins. Two additional factors beyond the choice of pigment are critical in the customization of the antennae to very different environments. These are the tailored structural organization of the pigments, and the tuning of pigment spectral properties by their in-vivo protein environment. In essence, all LHCs are composed of densely packed pigments that are usually encased in structure-preserving proteins and bound to membranes. The dense packing of pigments leads to strong electronic coupling between chromophores. Some LHCs also have a high degree of organization that aligns neighboring dipoles to further enhance electronic coupling. An example of this is the LH2 system found in purple bacteria, which consists of pigment-protein complexes in which the proteins form helical subunits enclosing rings of 18 and 9 pigments.³ This strong coupling, along with screening from solvent effects afforded by the binding to photosynthetic membranes, is believed to be the structural basis for the long-lived quantum coherent effects recently observed in a number of light harvesting complexes.⁴

Quantum mechanics also plays an important role in the performance of LHCs as antennae for light, both in determining their effective absorption cross-section and the excitation energy transfer subsequent to photon absorption. As mentioned above, the quantum mechanical coupling between multiple pigments can alter the oscillator strength of electronic transitions in pigment-protein complexes. It is generally believed that this feature helps to increase the efficiency of light absorption in several LHCs. The most striking example of this comes from green sulfur bacterium, a primitive photosynthetic organism that lives in extremely low light conditions. Green sulfur bacterium possesses a highly effective antenna structure, the chlorosome, that has recently been identified as large concentric nanotubes of

tightly packed bacteriochlorophyll molecules.⁵ This particular structure leads to very strong inter-pigment coupling and greatly enhanced electronic transition oscillator strengths for efficient light capture and energy transfer. Much of the drive for construction of artificial light harvesting complexes is to design and construct synthetic molecular complexes that mimic these features of the chlorosome.

The key to producing synthetic mimics of natural light harvesting systems is the establishment of the necessary distance relationships between multiple chromophores. Although this could, in principle, be achieved using elaborately designed synthetic molecules, this approach is typically quite laborious, is difficult to scale, and leads to highly aromatic systems with poor solubility and limited processing possibilities. A number of studies have instead used polymers and dendrimers as scaffold materials that establish an upper limit to the distance between chromophores,^{6,7} but these systems generally lack the rigidity needed to control transition dipole orientation and to prevent excimer-based quenching pathways.

As an alternative, several groups have developed ways to control the self-assembly of the chromophores themselves, generating large bundles of porphyrins that show energy transfer behavior.^{8,9} While these provide interesting chlorosome mimics, it is quite difficult to optimize the performance of these systems to meet specific applications, since the use of new chromophores with different optical properties can lead to unpredictable assembly outcomes. Alternatively, one can employ self-assembling protein coats of viruses as rigid scaffolds that can template the formation of synthetic light harvesting systems. In particular, architectures based on the capsid protein monomer of the tobacco mosaic virus (TMV) can be conveniently produced and employed for the assembly of chromophore arrays by introducing cysteine residues at specific positions that allow the covalent attachment of a wide variety of commercially-available chromophores with varied optical characteristics.¹⁰ One particularly interesting aspect of rod-like light harvesting arrays is the fact that they are inherently three-dimensional, and thus could possess redundant energy transfer pathways that could circumvent defect sites better than linear or ring-like systems.¹¹ An additional advantage of

this synthetic system is that the electronic properties of the aggregate complex can be chemically controlled by changing the type of chromophore, the type of linker used to covalently attach the chromophore to the protein,¹² and the position where it attaches to the protein.

In this work we investigate structural and spectroscopic features of synthetic pigment-protein structures for light harvesting that are based on TMV-templated chromophore assemblies. The close proximity of the chromophores in the TMV assemblies suggests that their excited electronic states will be closely coupled. To motivate the design and synthesis of new systems with enhanced electronic coupling, we analyze here several potential synthetic structures using theoretical modeling and spectroscopic characterization. We employ molecular mechanics simulations of the chromophore-protein systems to provide insight about the geometry and disorder. This is important given that these are systems for which crystal structures are hard to obtain, and thus direct experimental information about the geometry is lacking. A key focus of the present study is to understand both the geometry and the mobility of the chromophores, and the extent to which these factors are determined by the microscopic details of the surface of the protein, which typically forces the chromophores to fit into a solvent-accessible pocket. Different chromophores will be oriented differently and can have varying degrees of mobility depending on their point of linkage and the nature of the link to the protein. Such geometric and mobility information provides a systematic way to compare and screen for optimal chromophore-protein candidates for synthesis of artificial LHCs. The geometry of the chromophores is also critical to understanding the optical properties of these aggregate systems, since the electronic coupling between chromophores is primarily determined by the relative orientations of their transition dipole moments (TDMs).¹³ In the present work, the geometries of the conformers found from the molecular mechanics simulations are used in a tight-binding model to simulate the optical properties of the system, with comparison to experimental spectra.

The remainder of the paper is constructed as follows. Section 2 describes the TMV and chromophore structures employed here and summarizes the computational methods used

for the molecular mechanics structural studies with ground state chromophores, as well as the ab initio calculations for electronically excited chromophores and construction of the tight-binding model for simulation of the optical spectra. Section 3 presents the structural results with analysis of geometry, orientation and ordering of the chromophores, followed by presentation and analysis of the linear absorption spectra. Section 4 concludes with an assessment of the implications for computationally assisted molecular design of artificial light harvesting systems.

2 Computational Details

2.1 The TMV Protein and Chromophores

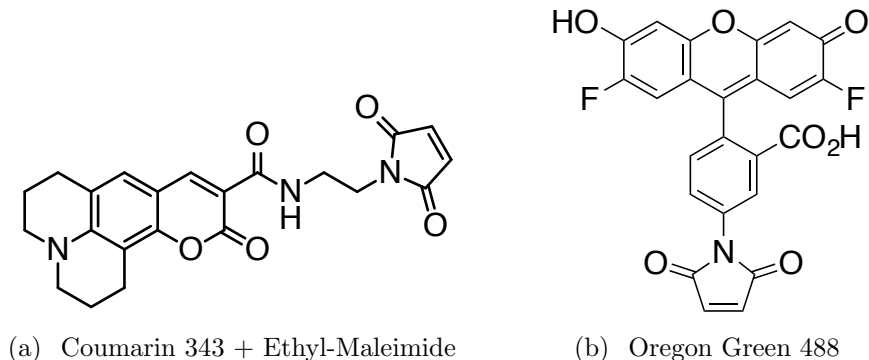


Figure 1: Molecular structures of modified Coumarin 343 (i.e., Coumarin 343 with a linker molecule, ethyl-maleimide) and Oregon Green 488.

The chromophore-protein complexes studied in this paper have all been experimentally synthesized.¹⁴ We have successfully attached chromophores to the TMV protein and the theoretical study of these complexes is the focus of this paper. The details of the self-assembly of TMV are available in ref. 10 and the details of attaching chromophores to the TMV protein are presented in ref. 14.

The TMV systems were self-assembled into a double-disk with 17-fold radial symmetry. We have studied the chromophores, Coumarin 343 and Oregon Green 488 (OG), attached

to the TMV protein at either the 104 (inner ring) or 123 (outer ring) residue positions. OG can be attached directly to the residues without modification while Coumarin 343 requires a linker molecule. Coumarin 343 was attached to both of the 104 and 123 positions with an ethyl-maleimide moiety. We refer to this complex of Coumarin 343 together with a linker molecule as CE. The molecular structures of CE and OG are available in Figure 1. The 104 and 123 positions differ in their distance from the center of the disk, thereby controlling the distance between neighboring chromophores, as illustrated in Figure 2. These systems will henceforth be referred to as CE-104, CE-123, OG-104, and OG-123.

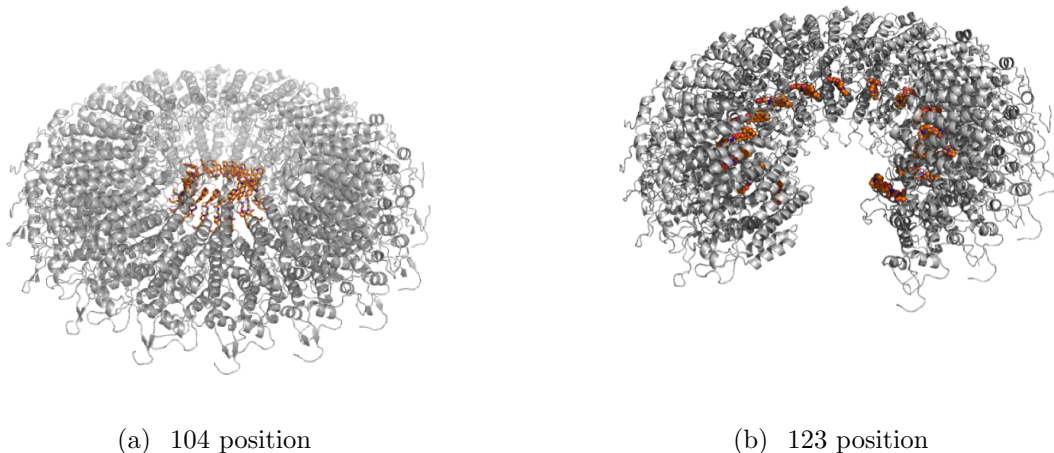


Figure 2: TMV-chromophore double-disk system. Protein colored grey, and chromophore colored orange. (a) Chromophore attached at the 104-residue position (inner ring) (b) Chromophore attached at the 123-residue position (outer ring)

2.2 Conformational search using Monte Carlo Multiple Minimum

The TMV-chromophore system is rather complex and is impossible to treat fully quantum-mechanically with today’s computational resources. Therefore, we explore the high-dimensional configurational space of the system using the Monte Carlo Multiple Minimum (MCMM) al-

gorithm with the force field of OPLS2005.¹⁵ All molecular mechanics simulations presented here were run with the Schrodinger’s MacroModel software suite.¹⁶ The double disk system has 34 monomers arranged with 17 monomers per layer, which includes approximately 200 rotatable bonds. For the simulations presented below, we focus on a single layer. An MC search over the full parameter space is computationally intractable, since the required time scales exponentially with the number of rotatable bonds.

For the CE-104 system we carried out computationally intensive simulations with the full 17 monomer ring system and compared this with simulations in which we considered only five monomers on each layer in the TMV under the assumption that chromophores that are separated by two or more monomers are non-interacting. We found that the average parameters (position and orientation) do not change significantly by using the truncated system. We therefore simulated only the truncated system for the other systems, CE-123, OG-104, and OG-123, and made the spectroscopic analysis on all four systems from the MCM configurations, using only the middle three chromophores on the top layer for the subsequent structural analyses.

The active space of the configurational search, are the atoms constituting the chromophore and linker molecules: these are represented by freely movable atoms. All protein atoms within 14 residues away from the chromophore constitute the first layer of nearby atoms. These are not freely movable, but are constrained to their initial position using a harmonic potential with a force constant of $200 \text{ kJ mol}^{-1} \text{ \AA}^{-2}$. All protein atoms between 15 and 29 residues away from the chromophore constitute the second layer of nearby atoms: these are effectively frozen in place using a stronger harmonic constraining potential. All protein atoms further away were ignored, as they are too distant to have any significant interaction with the free atoms. In the Monte Carlo molecular mechanics algorithm, given a starting molecule configuration, one randomly chooses some rotatable bonds (these are identified on the basis of the above specifications) and then rotates the atoms about these by a random angle. The potential energy of the resulting configuration is evaluated accord-

ing to the force field, and the configuration accepted if the value lies within an acceptable energy window. This yields new conformations of the molecule. The geometry of the new structure is then minimized to find a local minimum, using the Polak-Ribiere Conjugate Gradient (PRCG) method.¹⁷ For each configuration, we identify the coordinates of the chromophores and calculate the center of mass and the orientation of the TDM. To approximate the transition dipole moment for each chromophore, we pick three atoms on the perimeter of the conjugated rings of the chromophores to define a molecular plane. The normal vector to this plane is calculated for both the ground state DFT-optimized geometry and the MCMM geometry, and the rotation matrix that connects these two normal vectors is then applied to the TDM vector obtained from TDDFT (see Table 1). Since the conjugated rings generated by the MCMM simulations are quite rigid, validating the mapping of the MCMM configurations onto the DFT-optimized molecular structures, this is a cost-effective way to approximate the TDDFT TDM values of chromophores in the configurations generated by the MCMM calculations.

2.3 Ab-initio calculations of excited states

For the spectral simulations shown later, we need transition dipole moments (TDMs). This was achieved by performing time-dependent density functional theory (TDDFT)¹⁸ with B3LYP¹⁹/6-31G(d)^{20,21} for the DFT-optimized molecular geometries. Adiabatic excitation energies were extracted as the difference of TDDFT and DFT calculations at the DFT-optimized geometries. We truncated the linker molecule of CE and replaced the ethylmaleimide linker with a methyl group for simplicity. The TDDFT calculations employed 75 radial grid points and 302 Lebedev angular grid points. We also employed equation-of-motion coupled-cluster singles and doubles (EOM-CCSD)²² to further verify excitation energies and TDMs within the same basis set. These calculations were run with the development version of Q-Chem.²³

Table 1: Electronic structure results for the Coumarin-343 molecule with a methyl group linker substitute, showing excitation energies and corresponding ground-excited state transition dipoles and oscillator strengths, obtained from TDDFT-B3LYP, and EOM-CCSD calculations.

		TDDFT-B3LYP	EOM-CCSD
first excitation energy (eV)		3.4742	3.7753
transition dipole (a.u.)	x	2.6354	2.7087
	y	0.0057	0.1280
	z	0.0060	-0.0231
oscillator strength (a.u.)		0.5912	0.7609

In Table 1, we present the first excitation energy of the Coumarin-343 molecule with a methyl group linker substitute, obtained from the TDDFT-B3LYP and EOM-CCSD methods. We do not present the absolute energies of individual electronic states, since for TDDFT-B3LYP only the relative energies are meaningful, while the EOM-CCSD calculations are not converged with respect to the basis set size. In addition to the similar excitation energies and TDMs obtained from TDDFT-B3LYP and EOM-CCSD, the two methods also yield similar wavefunctions in terms of dominant electronic configurations. Since the first excited states in both chromophores are of singly-excited open-shell singlet character, the excitation energies from both of these two electronic structure methods are expected to be very accurate. The largest source of error is likely the limited size of basis set employed here, but in a previous paper²⁴ it was shown that employing a larger basis such as 6-311G* does not significantly affect the excitation energies (the change is 0.06 eV for both EOM-CCSD and TDDFT-B3LYP).

For the spectral simulations, we used the TDDFT-B3LYP energies and TDMs, rather than the corresponding EOM-CCSD values, for the following reasons. First, the TDDFT-B3LYP calculation is expected to be closer to its complete basis set limit, since the EOM-CCSD calculation is more sensitive to increasing the basis set size. Second, it was pointed by Koch et al.²⁵ that EOM-CCSD does not yield size-intensive TDMs, and thus EOM-CCSD may become a less reliable way to obtain TDMs for large systems. For these reasons, the analysis requiring TDMs (i.e., the spectral calculations) were all carried out using TDDFT-

B3LYP rather than EOM-CCSD.

2.4 Spectral Simulations

2.4.1 Hamiltonian Parametrization

A tight-binding Hamiltonian of the chromophores is often used to describe the electronic and optical properties of chromophore-protein systems^{26,27} and we employ a similar model here:

$$\hat{H} = \sum_{i=1}^N \epsilon_i \hat{a}_i^\dagger \hat{a}_i + \sum_{i=1}^N \sum_{j=1}^N J_{ij} \hat{a}_i^\dagger \hat{a}_j, \quad (1)$$

where ϵ_i is the on-site energy and J_{ij} is the coupling parameter between the site i and the site j . The coupling J_{ij} is a function of the positions of chromophores i and j , and of the relative orientations of each of their transition dipole moments. The close proximity between some of the chromophores in our system means that we cannot expect that the commonly used ideal dipole approximation (IDA)²⁸ for J_{ij} couplings will hold for all configurations.

We define $\Delta E_n^{\text{dimer}}$ to be the n -th electronic excitation energy of the dimer and $\Delta E_n^{\text{monomer}}$ to be the n -th electronic excitation energy of the monomer. In the case of a well-separated dimer, both $\Delta E_1^{\text{dimer}}$ and $\Delta E_2^{\text{dimer}}$ approach to $\Delta E_1^{\text{monomer}}$ and therefore the average of $\Delta E_1^{\text{dimer}}$ and $\Delta E_2^{\text{dimer}}$ is identical to $\Delta E_1^{\text{monomer}}$. However, our previous work shows that at close distances (< 12 Å), the average will begin to deviate from the monomer excitation energy ($\Delta E_1^{\text{monomer}}$). The extent of the deviation will depend on the distance and relative orientations of the two interacting chromophores.²⁴

In order to account for this geometry-dependent effect, we obtain more accurate Hamiltonian parameters based on pairwise TDDFT, as follows. The off-diagonal elements are given by

$$J_{ij} = J^{\text{TDDFT}}(\vec{R}_i, \vec{\mu}_i, \vec{R}_j, \vec{\mu}_j) = \frac{\Delta E_2^{\text{dimer}} - \Delta E_1^{\text{dimer}}}{2}, \quad (2)$$

and the diagonal elements are

$$\varepsilon_i = \Delta E_1^{\text{monomer}} + \Delta V_i \quad (3)$$

where

$$\Delta V_i = \sum_{j \neq i} V^{\text{TDDFT}}(\vec{R}_i, \vec{\mu}_i, \vec{R}_j, \vec{\mu}_j), \quad (4)$$

and

$$V^{\text{TDDFT}}(\vec{R}_i, \vec{\mu}_i, \vec{R}_j, \vec{\mu}_j) = \frac{\Delta E_2^{\text{dimer}} + \Delta E_1^{\text{dimer}}}{2} - \Delta E_1^{\text{monomer}}, \quad (5)$$

with \vec{R}_i the position of the i -th chromophore and $\vec{\mu}_i$ the orientation of the i -th chromophore.

In the full data set produced by the molecular mechanics configurations, there are about 10^7 pairwise interactions. Running a TDDFT calculation for each pairwise interaction in every configuration instance is computationally intractable and also redundant, since many of the pairs will have similar geometries. Our approach is therefore to instead run a TDDFT calculation at selected distances and orientations, and to interpolate between the values from these calculations to predict ab-initio values for other geometries. The precise procedure is as follows:

1. We parametrize the relative orientation of two interacting monomers: $r, \theta_A, \theta_B, \phi_B$ as defined in ref. 24.
2. We then discretize the space along those variables and calculate the TDDFT energies at the geometries defined by the following grid points: $r = [5 \text{ \AA}, 5.25 \text{ \AA}, 5.75 \text{ \AA}, 6 \text{ \AA}, 6.25 \text{ \AA}, 6.5 \text{ \AA}, 6.75 \text{ \AA}, 7 \text{ \AA}, 7.5 \text{ \AA}, 8 \text{ \AA}, 8.5 \text{ \AA}, 9 \text{ \AA}, 10 \text{ \AA}, 12 \text{ \AA}, 14 \text{ \AA}]$, $\theta_A = [-90^\circ, 90^\circ]$ with a 15° increment, $\theta_B = [0^\circ, 180^\circ]$ with a 15° increment, and $\phi_B = [0, 180]$ with 30° increment. Out of the possible 18928 geometries, we discard the points that yield unphysical geometries, which results in a training set of 4456 energies.
3. We employ a model function (*vide infra*) with three free parameters each for J_{ij} and V_{ij} . We fit these parameters by using linear regression to match J^{TDDFT} and V^{TDDFT} , respectively, at each of the grid points.

The model functions used to describe V and J are

$$J_{ij}^{\text{model}}(r, \theta_i, \theta_j, \phi_j) = C_J(r) J_{ij}^{\text{IDA}}(r, \theta_i, \theta_j, \phi_j) \quad (6)$$

$$V_{ij}^{\text{model}}(r, \theta_i, \theta_j, \phi_j) = C_V(r) V_{ij}^{\text{IDA}}(r, \theta_i, \theta_j, \phi_j) \quad (7)$$

$$C_k(r) = \frac{c_1^k}{(c_2^k - \exp(r/c_3^k))}, \quad k = J, V \quad (8)$$

where J_{ij}^{IDA} and V_{ij}^{IDA} are the couplings obtained from IDA, and C_k is a three-parameter logistical function whose value indicates whether the IDA brakes down. Specifically, if C_J and C_V deviate significantly from 1, then that is precisely when IDA breaks down. The fitted parameters are found to be: $c_1^V = 0.10388804$, $c_2^V = 0.21424357$, $c_3^V = 2.2014798$, $c_1^J = 0.2889955$, $c_2^J = 1.50871422$, $c_3^J = 3.14407206$.

Table 2 shows the average values of ΔV_i and J_{ij} obtained using this procedure for Coumarin 345 attached to TMV disks by the methyl group linker in an MCMM sampling of 10^5 molecular configurations.

Table 2: Effective Hamiltonian parameters for the Coumarin 345 and methyl linker system, averaged over all sites and 5000 MCMM geometry configurations on TMV double disks. V_{ij} are the diagonal Hamiltonian matrix elements (Eq. (3)), and J_{ij} are the off-diagonal Hamiltonian matrix elements (Eq. (2)).

	$\langle \Delta V_i \rangle$ [eV]	$\langle J_{ij} \rangle$ [eV]
CE-104	0.042	2.00×10^{-3}
CE-123	0.006	3.44×10^{-4}

2.4.2 Linear Absorption Spectra

In order to simulate the linear absorbance of the full double disk TMV system, we first sample small slices (i.e., 5 monomers) from the MCMM configurations and concatenate the middle three chromophores to generate the full system. Since 17 is a prime number, we need to take 5 samples of three chromophores, and 1 sample of two chromophores, all of which are sampled randomly. This is repeated for both the upper and lower disks. For each geometry

sample, we extract the center of mass positions and the transition dipole moments (R and μ) of each chromophore. Next, the tight binding Hamiltonian in Eq. (1) is constructed using the parameters ε_i and J_{ij} described in Eq. (6). The Hamiltonian is diagonalized by a unitary transformation to yield exciton states and energies:

$$\hat{H}|\psi_k\rangle = E_k|\psi_k\rangle \quad (9)$$

$$|\psi_k\rangle = \sum_i^N c_{ik}|\phi_i\rangle. \quad (10)$$

The linear absorption spectrum for a given Hamiltonian is then calculated using

$$\vec{\mu}_k = \sum_i^N c_{ik}\vec{\mu}_i \quad (11)$$

$$\text{Abs}(E) \propto \sum_k^N \|\vec{\mu}_k\|^2 \exp \left[-\frac{(E - E_k)^2}{2\sigma^2} \right], \quad (12)$$

where $\vec{\mu}_k$ is the transition dipole moment of exciton k , obtained by transforming the vector of molecular transition dipole moments with the same unitary transformation that diagonalizes the Hamiltonian. The summation in Eq.(12) describes a discrete convolution between a gaussian function, and a stick spectrum composed of excitations from the ground to single excitation states of the Hamiltonian, with weights given by the 2-norm squared of the exciton transition dipole moment. The variance of the gaussian function (σ) is the line broadening parameter for our simulated linear absorption spectrum at a given geometry configuration and corresponds to the homogeneous linewidth. In the limit where $\sigma \rightarrow 0$, we recover the eigenvalue stick spectrum. Eq. (12) yields the spectrum for a single geometry instance. This process is repeated 5000 times (large enough to obtain converged spectra) to average over the different possible geometry configurations, and is then normalized by the maximum absorbance.

3 Results and Discussion

3.1 Geometric Distributions of Chromophores

We analyze the MCMC conformations based on the center of mass (CM) positions of the chromophores and the orientations of TDMs of the chromophores. Those two collective variables are particularly useful in understanding the geometric distribution, as we shall see below.

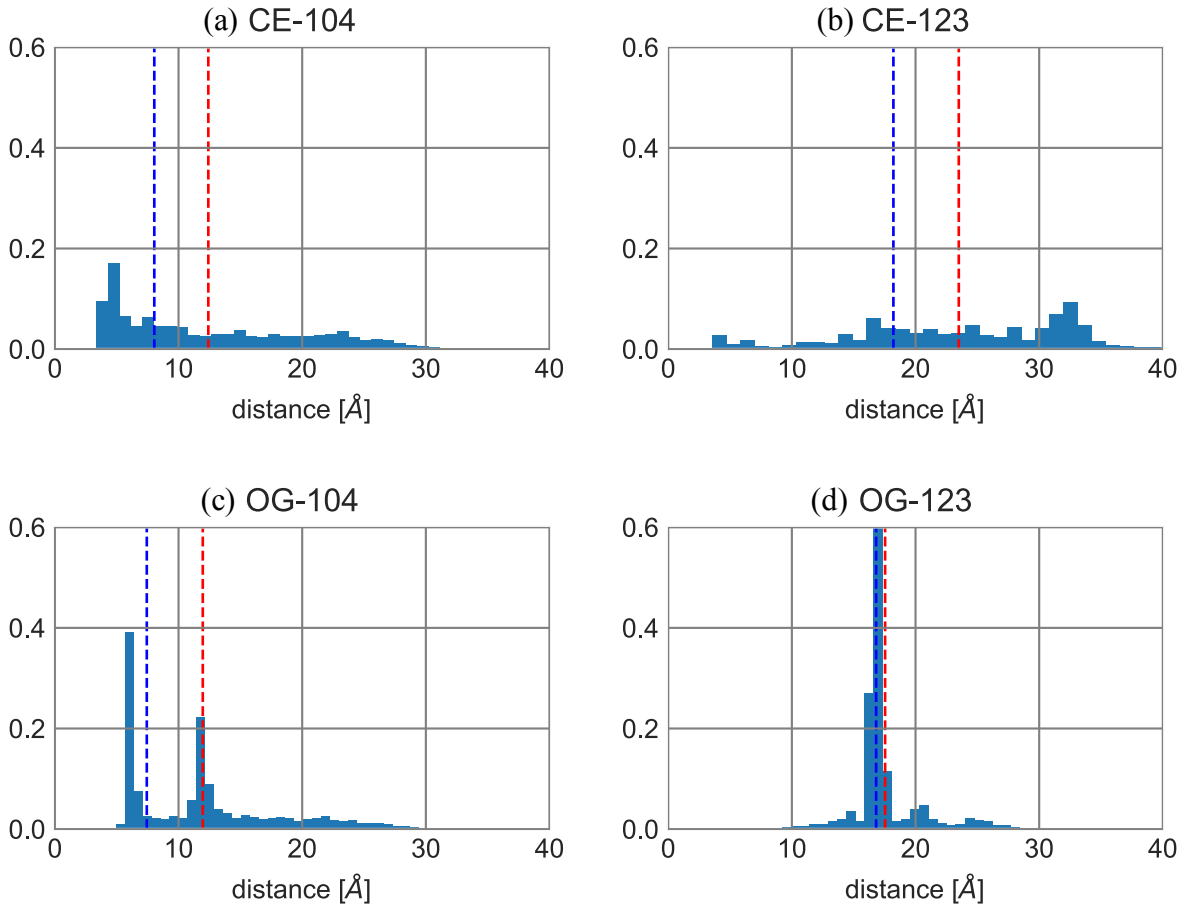


Figure 3: Normalized histograms of the center-of-mass distances between nearest-neighbor chromophores in (a) CE-104, (b) CE-123, (c) OG-104, and (d) OG-123. The blue dotted line in each panel indicates the “ideal” nearest-neighbor distance, r_{ideal} , and the red dotted line indicates the mean distance, \bar{r} (see text). The number of samples used in each histogram is 3607, 5499, 5918, and 5810, respectively.

Figure 3 shows a histogram of distances between the CM positions of nearest-neighbor

chromophores. We first note that both CE-104 and CE-123 exhibit significant multimodal behaviors, while bimodal and monomodal behaviors are observed for OG-104 and OG-123, respectively. The qualitative difference between CE and OG can be explained simply: the linker molecule in CE allows Coumarin to move easily its CM position whereas OG has no linker molecule in our study. We further computed the “ideal” nearest-neighbor distance, r_{ideal} , which assumes an equilateral 17-polygon, and the mean of nearest-neighbor distances, \bar{r} . Those two values are not meaningful in the case of highly multimodal histograms as in the CE cases. In the case of OG-123, two values are almost identical, whereas each of two peaks in OG-104 roughly corresponds to r_{ideal} and \bar{r} .

The significance of Figure 3 is that some chromophores (in particular CE-104 and OG-104) in the TMV systems are not far enough apart for the IDA to be valid; the distance between nearest neighbors is often less than 12 Å. Based on our previous study of Coumarin 343, when two chromophores are closer than 12 Å, it is likely that IDA to the Hamiltonian starts to fail quite catastrophically.²⁴ This was indeed our motivation to go beyond IDA, and this will be discussed further below. We note that in the previous study,²⁴ Coumarin was considered without a linker molecule. However, we expect the failure of dipole approximations to appear similarly for Coumarin attached to a linker molecule. We can also expect the same behavior for the OG chromophore.

For the purpose of analysis, we introduce the monomer frame illustrated in Figure 4. We define the monomer frame as follows: for every chromophore, the x-axis points towards the center of TMV disk, the z-axis is parallel to the axis of rotational symmetry, and the y-axis is defined in the conventional way given these x- and z-axes (i.e., $\hat{y} = \hat{z} \times \hat{x}$). The radial axis of the polar plot ranges from 0° to 180° and corresponds to the polar angle of the monomer frame. The angular axis of the polar plot ranges from 0° to 360° and corresponds to the azimuthal angle of the monomer frame.

Figure 5 shows a histogram of the orientation of TDMs measured in the monomer frame. It shows that the TDMs for the CE systems have broader angular distributions than the

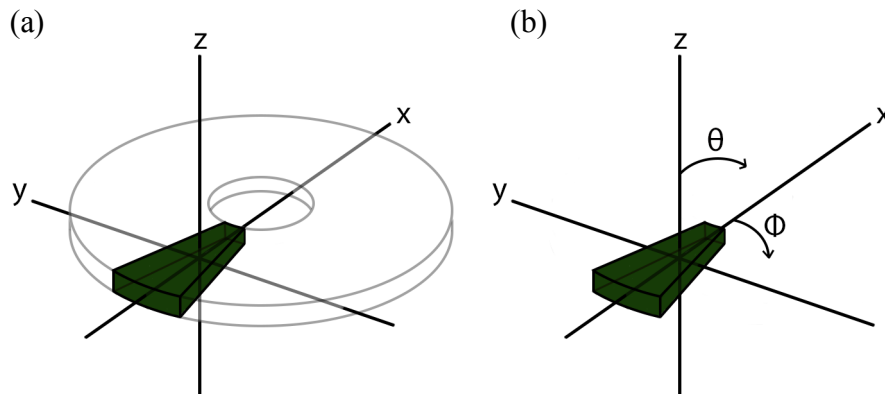


Figure 4: (a) Schematic description of the monomer frame embedded into the entire TMV represented by a disk. The monomer is represented by the wedge. (b) Schematic description that represents the polar angle θ and the azimuthal angle ϕ in the monomer frame.

TDMS in the OG systems. In other words, OG systems are far more confined in their orientation than CE systems. This does not necessarily mean that OG systems are more ordered. The quantification of order-disorder will be discussed later in the paper. Instead, the wider distribution of the TDMs in the CE systems can be understood by recognizing the effect of the presence of a linker molecule. We see similar trends in both chromophores when attached to the 104 or 123 position. The 104 position exhibits a wider vertical spread compared to the 123 position, whereas the 123 position shows a wider horizontal spread compared to the 104 one. The broad distribution of the CE systems is somewhat surprising, given that the chromophore molecules are surrounded by the TMV protein environment. However, the linker molecule gives enough flexibility to the Coumarin chromophore which results in a broad distribution of geometries. This is one of the reasons that make atomistic simulations of the system challenging and computationally expensive.

3.2 Order, Disorder, and Correlation Among Chromophores

We now discuss one-body and two-body observables that can be extracted from the chromophore distributions, in order to further quantify the balance between order and disorder present in the chromophore-TMV system. There is a simple analogy between our system

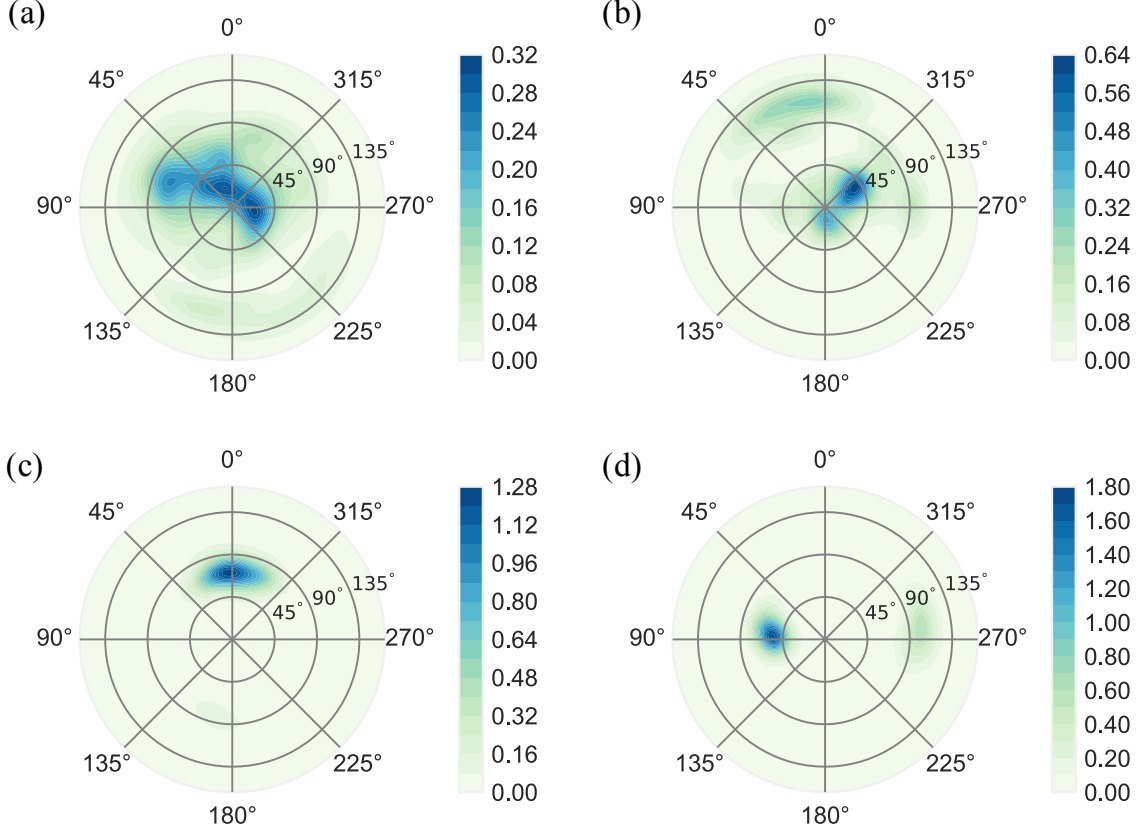


Figure 5: Normalized distributions of the transition dipole moments orientation for (a) CE-104, (b) CE-123, (c) OG-104, and (d) OG-123. The radial distance indicates θ and the angular orientation indicates ϕ . θ and ϕ are defined in Figure 4. We fitted the TDM vectors to a bivariate gaussian kernel density estimator with a 0.2 bandwidth to obtain this figure.

and one-dimensional classical Heisenberg model of 17 sites with periodic boundary conditions. In other words, a TMV disk can be reduced down to a lattice with 17 sites and the orientation of each of the 17 chromophores can be considered analogous to a classical spin on the corresponding site. This analogy allows us to utilize one-body and two-body measures that are widely used to quantify order in spin systems. In passing we note that the orientation vectors of TDMs used in the following analyses are all normalized and measured in the monomer frame. As we analyze only five monomers per sample, periodic boundary conditions were not applied here.

The one-body observable considered here is the average of the magnetization of spins,

Table 3: The average of the TDM vector orientation (a.u.) along each cartesian axis in the monomer frame for the CE and OG systems.

	CE-104	CE-123	OG-104	OG-123
$\langle \mu_x \rangle$	0.336	0.179	0.587	0.069
$\langle \mu_y \rangle$	-0.253	0.220	-0.058	-0.194
$\langle \mu_z \rangle$	0.410	0.458	0.148	0.145

Table 4: The average of the absolute value of the TDM vector (a.u.) along each cartesian axis in the monomer frame for the CE and OG systems.

	CE-104	CE-123	OG-104	OG-123
$\langle \mu_x \rangle$	0.558	0.389	0.824	0.159
$\langle \mu_y \rangle$	0.408	0.464	0.284	0.821
$\langle \mu_z \rangle$	0.538	0.609	0.293	0.429

which in our case is the average of the orientation vector of TDMs, defined as

$$\langle \vec{\mu} \rangle = \frac{1}{N_{\text{samples}} N_{\text{spins}}} \sum_i^{N_{\text{samples}}} \sum_{\alpha}^{N_{\text{spins}}} \vec{\mu}_{\alpha}(i). \quad (13)$$

Here $\langle \vec{\mu} \rangle$ is a normalized TDM vector and the magnitude of each cartesian component of $\vec{\mu}_{\alpha}(i)$ ranges from 0 to 1. In the case of ferromagnets, this measure is sufficient to determine whether the system is ordered. A small value of $\vec{\mu}$ indicates a disordered phase and a large value indicates an ordered phase. However, in the case of antiferromagnets, a small value of $\vec{\mu}$ is not enough to conclude that it is a disordered phase. This is because a perfect antiferromagnet would exhibit negligible average magnetizations.

Table 3 shows the averages of the projections of the TDM orientation vectors along each cartesian axis in the monomer frame. CE and OG present a qualitative difference here, since OG has at least one direction with a very small value. The small values along the y -axis in OG-104 and the x -axis in OG-123 are particularly interesting, since they may indicate an antiferromagnetic ordering along those axes. To further investigate this, we computed $\langle |\vec{\mu}| \rangle$ which is defined similarly to Eq. (13), with $|\vec{\mu}_{\alpha}(i)|$ replacing $\vec{\mu}_{\alpha}(i)$. These results are presented in Table 4. If there is no difference between $\langle |\mu| \rangle$ and $\langle \mu \rangle$ then the system is either ferromagnetic or unpolarized, while a significant difference between them suggests an

antiferromagnetic ordering. Both the y -component of OG-104 and the x -component of OG-123 do show a significant difference, which could therefore be taken as indicating possible antiferromagnetism along those axes. However, analysis of the two-body correlations below will show that these two components have no ordering.

We have investigated a two-body correlation function, (i.e., a two-point correlator), namely the spin-spin correlation function, C_{spin} ,

$$C_{\text{spin}} = \sum_{\langle\alpha\beta\rangle} \langle \vec{\mu}_\alpha \cdot \vec{\mu}_\beta \rangle = \frac{1}{N_{\text{samples}} N_{\text{neighbors}}} \sum_i^{N_{\text{samples}}} \left(\sum_{\langle\alpha\beta\rangle} \vec{\mu}_\alpha^i \cdot \vec{\mu}_\beta^i \right). \quad (14)$$

Here the sum goes over nearest neighbors α, β . We have included only nearest neighbor correlations in C_{spin} , even though the underlying interaction between spins in our case is long-ranged. This was done on purpose because the interaction is dominated by nearest neighbor interactions and this truncation leads to a simple interpretation of the physical meaning of C_{spin} . In particular, the values of Eq. (14) range between -1 and 1, where the limit of 1 corresponds to perfect ferromagnetic order, -1 corresponds to perfect antiferromagnetic order, and 0 indicates either no order, i.e., perfect disorder. We also define C_{spin}^γ with $\gamma \in \{x, y, z\}$ to quantify these same types of orders along each cartesian axis, with each component defined similarly to Eq. (14), but using TDM components $[\mu_\gamma]_\alpha^i$.

Table 5: Spin-spin correlation functions (Eq. (14)) for each system

	CE-104	CE-123	OG-104	OG-123
C_{spin}	0.433	0.390	0.386	0.035
C_{spin}^x	0.108	0.078	0.376	0.006
C_{spin}^y	0.101	0.026	-0.007	0.018
C_{spin}^z	0.224	0.286	0.017	0.011

Table 5 shows the values of this two-body observable. We see that CE-104, CE-123 and OG-104 are all considerably more ordered than OG-123. Both CE-104 and CE-123 show partial ferromagnetic ordering. This is strongly anisotropic in the case of CE-123 and weakly anisotropic in the case of CE-104, with greater ferromagnetic correlations along the z -axis

perpendicular to the plane of the disk in both cases. In contrast, Table 4 implies that CE-123 has non-negligible orientation along all three x, y, z -axes. Taken together with Table 5, this suggests that CE-123 has weakly correlated partial ferromagnetic order along the x, y -axes and more strongly correlated ferromagnetic order that is consequently also greater in extent along the z -axis. Similarly, Tables 4 and 5 imply that OG-104 exhibits ferromagnetic ordering along x -axis and disorders along the y, z -axis. As illustrated in Figure 5, OG-104 is also strongly confined around the positive x -axis. Therefore, OG-104 is confined and at the same time well-ordered. OG-104 would thus be a good future candidate for further theoretical studies, since the high degree of both spatial confinement and chromophore ordering means that the entire conformation space need not be explored. OG-123 is interesting in the sense that Table 5 shows it is disordered along every axis. However, according to Figure 5, it is nevertheless confined in space. Although OG-123 is spatially confined by the TMV protein environment, the relative orientation between different chromophores is almost completely random.

3.3 Linear Absorption Spectra of Coumarin-TMV double disks

Figure 6 shows simulated linear absorption spectra for the Coumarin TMV-templated aggregates CE-104 and CE-123, and compares these with the corresponding experimental linear absorbance spectra for the fully loaded disks containing 17 chromophores on each side of the disk.¹⁴ The baseline has been subtracted from the experimental spectra to remove contributions of light scattering. The left panel, Figure 6(a), shows the simulated spectra resulting from averaging over the MCMM geometries, using the TDDFT-derived Hamiltonian parameters in Eq. (6) for the CE-104 and CE-123 system as described in Section 2.4. The right panel, Figure 6(b), shows simulated spectra for an idealized 17-fold symmetric structure on a single disk, which has C_{17h} symmetry (see Fig. 7 and text). In all of these Coumarin 343 spectral simulations, we set $\Delta E_1^{\text{monomer}} = 2.8$ eV (442.8 nm), which results in MCMM-averaged diagonal Hamiltonian energies $\varepsilon_i = 2.842$ eV and 2.806 eV for CE-104 and CE-123,

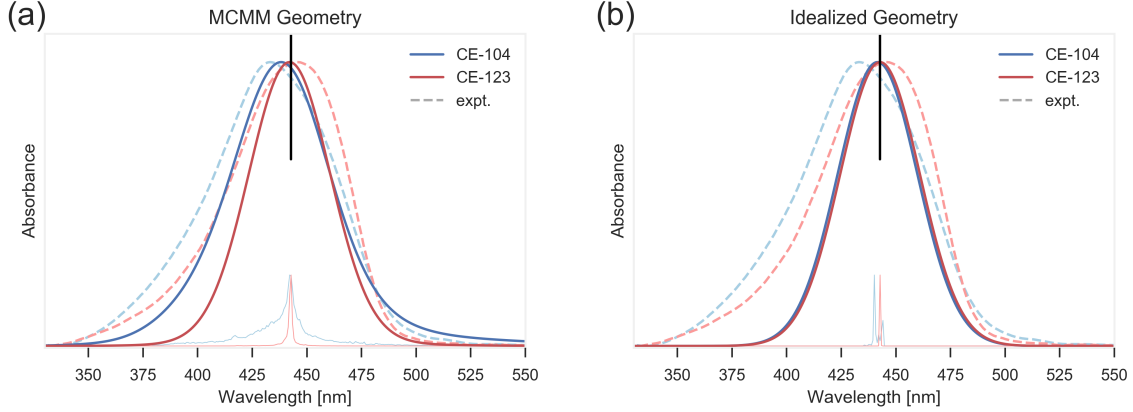


Figure 6: Simulations of linear absorption spectra using the TDDFT-improved Hamiltonian parameters in Eq. 6. (a) Spectra averaged over the MCMC-geometries (b) spectra of a single idealized, 17-fold symmetric geometry (see Fig. 7). Dotted lines are the absorption spectra obtained from experiment, and the vertical black line is the monomer excitation energy $\Delta E^{\text{monomer}}$. The line broadening parameter σ is set to 0.12 eV (36.4 nm at a peak position 442.8 nm corresponding to the shifted monomer energy $\Delta E_1^{\text{monomer}} = 2.8$ eV (see text). Below these simulated spectra, we show simulations of the same system using 0.001 eV (0.32 nm at a peak position 442.8 nm).

respectively. The $\Delta E_1^{\text{monomer}}$ energy is considerably smaller than the vacuum TDDFT excitation energy for the CE system (3.47 eV) reported in Table 1, and was shifted from this in order to achieve an optimal average coincidence of the peaks of the simulated spectra with those of the experimental spectra. This results in a global energy shift to the diagonal entries of the Hamiltonian that accounts for the interaction of the Coumarin 343 with the solvent, and the reorganization energy of the chromophores when attached to TMV. The choice of $\Delta E_1^{\text{monomer}}$ does not affect our analysis of the difference between the absorption peak locations for CE-104 and CE-123, since our conclusions are all based on relative energy differences. The line broadening parameter σ is set to 0.12 eV (36.4 nm at a peak position 2.8 eV \equiv 442.8 nm) for all simulated spectra. This parameter provides a qualitative fit to the overall spectral distribution and represents implicit dependency on both the homogeneous line broadening that results from coupling to vibrational degrees of freedom and the inhomogeneous line broadening arising from static disorder in the transition energies.

We note that the linewidth parameter ~ 36 nm is the same order of magnitude as

that of typical photosynthetic chromophores in solution¹ and just as for aggregates of such chromophores in light harvesting, it is larger than the inter-chromophore coupling J (see Table 2). At the bottom of each of these plots, we also show the same spectral simulations with σ set to 0.001 eV (0.32 nm at a peak position 442.8 nm). Since this broadening value is lower than the resolution of the x-axis, these insert plots can be interpreted as histograms of the stick spectra from ground to discrete excited state energies, which corresponds to heterogeneous line broadening.

We first discuss the simulated spectra resulting from averaging over the MCMC geometries, panel (a), and then the spectra resulting from averaging over instantaneous idealized C_{17h} symmetric geometries, panel (b).

Figure 6(a) shows that the distribution of the lines in the stick spectrum for the CE-123 system is considerably narrower than that for the CE-104 system. This is due to the fact that in all geometric configurations of the CE-123 system the chromophores are further apart than in the CE-104 system, resulting in smaller off-diagonal couplings and a consequent smaller range of the electronic excited state eigenvalues. In the limit of infinite separation, these excited state eigenvalues become degenerate and would yield a delta function spectrum. In comparison, the CE-104 system (blue line) has a much wider stick spectrum, showing non-zero absorption between 400-475 nm. This is due in part to the greater range of the electronic excited state eigenvalues for CE-104 (~ 2.22 eV), and also to the greater contribution of static disorder in the chromophore positions and orientations (the range of individual excited state eigenvalues for CE-104 averaged over all geometry instances is 0.80 eV). Both of these factors contribute to the significant inhomogeneous broadening of the CE-104 spectral absorption. It is also apparent that the stick spectrum distribution of CE-104 shows a significant asymmetrical weight towards higher energies, while that of CE-123 is more symmetrical. This results in a greater asymmetry in the convolved spectrum for CE-104, consistent with the greater broadening of the experimental spectrum on the blue side of the peak. This greater asymmetry for CE-104 also contributes to the overall blue

shift of the CE-104 relative to the CE-123 system.

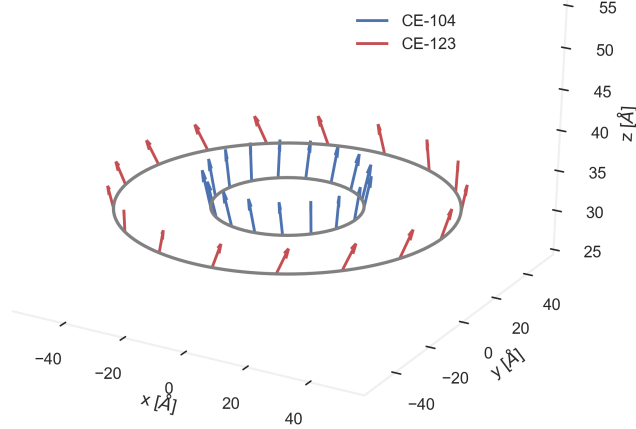


Figure 7: An idealized C_{17h} symmetric geometry where the positions and orientations of the chromophores are obtained from the average over all the MCMM configurations. Arrows indicate the orientation of the transition dipole moment. The blue arrows describe the CE-104 system, while the red arrows describe the CE-123 system.

To obtain more insight into the effect of the chromophore geometries on the absorption line shape, Figure 6(b) shows a computed spectrum for an ideal C_{17h} symmetric geometry that is constructed by setting the positions and orientations of the chromophores to the average values over all the MCMM configurations. Here the spread in the stick spectrum now derives solely from the excitonic splitting of the 17 excited state energies, since all chromophores have the same position and orientation on the TMV disk. The resulting idealized geometry on the top half of a double disk is shown for both the CE-104 and CE-123 systems in Figure 7. The symmetry of these geometries implies that every nearest-neighbor pair of chromophores has an identical geometry relative to each other, so no pair couples more strongly than any of the others. For such idealized geometries, the resulting Hamiltonian will yield eigenvalues with a smaller spread than the range of values obtained from sampling with MCMM. Consequently, we see a tighter spectrum for CE-104 in Figure 6(b) than in

Figure 6(a).

Table 6: Mean and Maximum Wavelength of Absorbance. Δ_{avg} is the difference between CE-104 and CE-123 mean wavelengths ($\lambda_{\text{CE-123}}^{\text{avg}} - \lambda_{\text{CE-104}}^{\text{avg}}$), and Δ_{max} is the difference between CE-104 and CE-123 peaks ($\lambda_{\text{CE-123}}^{\text{max}} - \lambda_{\text{CE-104}}^{\text{max}}$).

	$\lambda_{\text{CE-104}}^{\text{avg}}$ [nm]	$\lambda_{\text{CE-123}}^{\text{avg}}$ [nm]	Δ_{avg} [nm]	$\lambda_{\text{CE-104}}^{\text{max}}$ [nm]	$\lambda_{\text{CE-123}}^{\text{max}}$ [nm]	Δ_{max} [nm]
Experiment ¹⁴	429.28	435.14	5.86	433.07	445.07	12.00
MCMM	438.32	442.78	4.46	437.46	442.03	4.57
Idealized	442.42	443.60	1.18	445.13	443.06	-2.06

We quantify the differences between the CE-104 and CE-123 spectra in Figure 6 by calculating the peak and mean wavelength of absorption for each spectrum, where the latter is given by the weighted average:

$$\lambda^{\text{avg}} = \frac{\sum_{\lambda} \text{Abs}(\lambda) * \lambda}{\sum_{\lambda} \text{Abs}(\lambda)}. \quad (15)$$

The values are shown in Table 6 for both the MCMM averaged spectra and the idealized spectrum, where they are also compared with the peak absorption values obtained from experimental spectra.¹⁴ The peaks in the experimental spectrum exhibit a blue shift when going from the CE-123 system to the CE-104 system. Taken together with the narrow distribution of the individual excitonic absorptions in the CE-123, this is consistent with the lack of any significant interaction-induced shift of the chromophores in the CE-123 system, while the more closely packed ring of aggregates in the CE-104 system behaves as an H-aggregate. Table 6 shows that the simulation using the MCMM-derived geometries is better able to reproduce this relative blue shift of the experimental spectra than the simulation using the idealized 17-fold structure. While both the simulations using the MCMM geometries and the idealized geometries exhibit a blue shift, the simulation that incorporates disorder is able to match this trend much better. We conclude that the disorder in the geometry of the CE-104 and CE-123 systems is an important feature of these systems, and it is important to account for this disorder in order to model the optical properties of TMV-templated chromophore aggregates accurately.

4 Conclusions

In this work, we present a protocol that generates conformations using a Monte Carlo multiple minima (MCMM) conformation search algorithm, parametrizes a semiempirical tight-binding Hamiltonian based on ab-initio TDDFT calculations, and combines these to generate a linear absorption spectrum that can be directly compared to experiments.

We applied this protocol to study a recently synthesized artificial light harvesting system consisting of chromophores attached to a tobacco mosaic virus (TMV) protein. We studied Coumarin 343 together with a linker, and Oregon Green 488, both of which were attached to the 104 and 123 sites on the TMV protein. The resulting four systems, CE-104, CE-123, OG-104, and OG-123, were studied with MCMM and we obtained a wide array of local minima. We characterized those conformers using the orientation of the transition dipole moment and center-of-mass of dyes attached to the TMV protein. Such a characterization led to a qualitative and quantitative understanding of structural order and disorder associated with the dyes. CE-104 and CE-123 both exhibit a very broad geometric distribution, which makes any more detailed theoretical study relatively intractable. OG-104 and OG-123 are relatively spatially well confined, but OG-123 is more disordered than is OG-104 in terms of the spin-spin correlation function discussed in the main text. For future studies, we therefore conclude that OG-104 will likely be the system most suited for more detailed theoretical study.

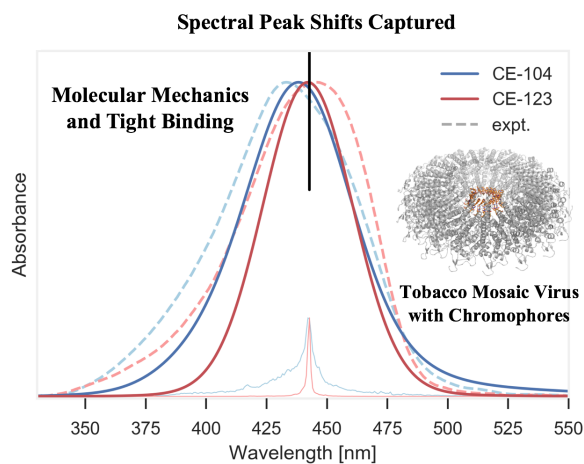
Lastly, we combined the wide array of conformations found through MCMM with a semiempirical tight-binding Hamiltonian for the Coumarin-linker system to calculate linear absorption spectra of CE-104 and CE-123, and compared this with experimental spectra. We observed that it is necessary to account for the proper distribution over geometries of conformations to properly reproduce the experimentally observed blue shift of the CE-104 system relative to the CE-123 system. We also confirmed the greater asymmetry of the lineshape of the CE-104, with detailed analysis showing that this derives largely from the greater distribution of the excitonic energies for CE-104, due to the closer distance between chromophores. It is encouraging that a qualitatively accurate spectrum could be obtained

in-silico from MCMM calculations together with a tight-binding Hamiltonian. A challenge for further work is to incorporate the interaction with vibrational degrees of freedom, as well as to develop more reliable and economical ways to generate energy minima and thereby increase the conformational sampling..

5 Acknowledgement

This work was supported by the DARPA QuBE (Quantum Effects in Biological Environments) program under contract number N66001-10-1-4068. The views expressed are those of the authors and do not reflect the official policy or position of the Department of Defense or the US Government. J. L. thanks POSTECH for the exchange student scholarship that made early stages of this work possible.

6 Table of Contents Graphic



References

- (1) Blankenship, R. E. *Molecular Mechanisms of Photosynthesis*, 2nd ed.; Chichester, UK: Wiley-Blackwell, 2014.

- (2) Van Amerongen, H.; Valkunas, L.; Van Grondelle, R. *Photosynthetic Excitons*; World Scientific, 2000.
- (3) McDermott, G.; Prince, S.; Freer, A.; Hawthornthwaite-Lawless, A.; Papiz, M.; Cogdell, R.; Isaacs, N. Crystal Structure of an Integral Membrane Light-Harvesting Complex from Photosynthetic Bacteria. *Nature* **1995**, *374*, 517–521.
- (4) Scholes, G. D.; Fleming, G. R.; Chen, L. X.; Aspuru-Guzik, A.; Buchleitner, A.; Coker, D. F.; Engel, G. S.; van Grondelle, R.; Ishizaki, A.; Jonas, D. M. et al. Using Coherence to Enhance Function in Chemical and Biophysical Systems. *Nature* **2017**, *543*, 647–656.
- (5) Ganapathy, S.; Oostergetel, G. T.; Wawrzyniak, P. K.; Reus, M.; Chew, A. G. M.; Buda, F.; Boekema, E. J.; Bryant, D. A.; Holzwarth, A. R.; de Groot, H. J. Alternating Syn-Anti Bacteriochlorophylls form Concentric Helical Nanotubes in Chlorosomes. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 8525–8530.
- (6) Gilat, S. L.; Adronov, A.; Frechet, J. M. Light Harvesting and Energy Transfer in Novel Convergently Constructed Dendrimers. *Angew. Chem. Int. Ed.* **1999**, *38*, 1422–1427.
- (7) Kim, J.; McQuade, D. T.; Rose, A.; Zhu, Z.; Swager, T. M. Directing Energy Transfer Within Conjugated Polymer Thin Films. *J. Am. Chem. Soc.* **2001**, *123*, 11488–11489.
- (8) Choi, M.-S.; Yamazaki, T.; Yamazaki, I.; Aida, T. Bioinspired Molecular Design of Light-Harvesting Multiporphyrin Arrays. *Angew. Chem. Int. Ed.* **2004**, *43*, 150–158.
- (9) Wang, Z.; Medforth, C. J.; Shelnutt, J. A. Porphyrin Nanotubes by Ionic Self-Assembly. *J. Am. Chem. Soc.* **2004**, *126*, 15954–15955.
- (10) Miller, R. A.; Presley, A. D.; Francis, M. B. Self-Assembling Light-Harvesting Systems from Synthetically Modified Tobacco Mosaic Virus Coat Proteins. *J. Am. Chem. Soc.* **2007**, *129*, 3104–3109.

- (11) Miller, R. A.; Stephanopoulos, N.; McFarland, J. M.; Rosko, A. S.; Geissler, P. L.; Francis, M. B. Impact of Assembly State on the Defect Tolerance of TMV-Based Light Harvesting Arrays. *J. Am. Chem. Soc.* **2010**, *132*, 6068–74.
- (12) Delor, M.; Dai, J.; Roberts, T. D.; Rogers, J. R.; Hamed, S. M.; Neaton, J. B.; Geissler, P. L.; Francis, M. B.; Ginsberg, N. S. Exploiting Chromophore-Protein Interactions through Linker Engineering To Tune Photoinduced Dynamics in a Biomimetic Light-Harvesting Platform. *J. Am. Chem. Soc.* **2018**, *140*, 6278–6287.
- (13) Scholes, G. D.; Fleming, G. R.; Olaya-Castro, A.; van Grondelle, R. Lessons from Nature About Solar Light Harvesting. *Nat. Chem.* **2011**, *3*, 763–774.
- (14) Finley, D. Synthesis and Development of Viral Capsid Templated Light Harvesting Systems. PhD, UC Berkeley, 2014.
- (15) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.
- (16) Schrödinger Release 2012: MacroModel, Schrödinger, LLC, New York, NY, 2012.
- (17) Polak, E.; Ribiere, G. Note sur la Convergence de Méthodes de Directions Conjuguées. *Esaim Math. Model. Numer. Anal.* **1969**,
- (18) Dreuw, A.; Head-Gordon, M. Single-Reference Ab Initio Methods for the Calculation of Excited States of Large Molecules. *Chem. Rev.* **2005**, *105*, 4009–4037.
- (19) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (20) Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. A Complete Basis Set Model Chemistry. I. The Total Energies of Closed-

- shell Atoms and Hydrides of The First-Row Elements. *J. Chem. Phys.* **1988**, *89*, 2193–2218.
- (21) Petersson, G. A.; Al-Laham, M. A. A Complete Basis Set Model Chemistry. II. Open-Shell Systems and The Total Energies of The First-Row Atoms. *J. Chem. Phys.* **1991**, *94*, 6081–6090.
- (22) Krylov, A. I. Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhiker’s Guide to Fock Space. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433–462.
- (23) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X. et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.
- (24) Lee, D.; Greenman, L.; Sarovar, M.; Whaley, K. B. Ab Initio Calculation of Molecular Aggregation Effects: A Coumarin-343 Case Study. *J. Phys. Chem. A* **2013**, *117*, 11072–11085.
- (25) Koch, H.; Kobayashi, R.; De Merás, A. S.; Jørgensen, P. Calculation of Size-Intensive Transition Moments from the Coupled Cluster Singles and Doubles Linear Response Function. *J. Chem. Phys.* **1994**, *100*, 4393–4400.
- (26) van Amerongen, H.; van Grondelle, R.; Valkunas, L. *Photosynthetic Excitons*; World Scientific, 2000.
- (27) Scholes, G. D.; Rumbles, G. Excitons in Nanoscale Systems. *Nat. Mater.* **2006**, *5*, 683–696.
- (28) Kasha, M.; Rawls, H. R.; El-Bayoumi, M. A. The Exciton Model in Molecular Spectroscopy. *Pure and Applied Chemistry* **1965**, *11*.